

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 18, 2006, 00:41:40 ; Search time 4618 Seconds  
(without alignments)  
7156.415 Million cell updates/sec

Title: US-10-733-816-2  
Perfect score: 2081  
Sequence: 1 MEYXPMGEGMGSRPTTSF.....QELSSNPPLATILIPHARI 394

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues  
Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framet p2n.model -DEV=exp  
-Q=/abs/ABSSWEB\_spool/US10733816/runat\_15092006.085446.15519/app.query.fasta\_1  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03p  
-USER=US10733816 ACGN.1.1.9528 @runat\_15092006.085446.15519 -NCPU=6 -ICPU=3  
-NO\_MAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_est1:  
2: gb\_est3:  
3: gb\_est4:  
4: gb\_est5:  
5: gb\_est6:  
6: gb\_htc:  
7: gb\_est2:  
8: gb\_est7:  
9: gb\_est8:  
10: gb\_est9:  
11: gb\_gsl1:  
12: gb\_gsl2:  
13: gb\_gsl3:  
14: gb\_gsl4:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID       | Description        |
|------------|--------|-------------|--------|-------------|--------------------|
| 1          | 2015   | 96.8        | 2312   | 6 AK170965  | AK170965 Mus muscu |
| 2          | 2015   | 96.8        | 2847   | 6 AK154293  | AK154293 Mus muscu |
| 3          | 2013   | 96.7        | 1732   | 6 CR617019  | CR617019 full-leng |
| 4          | 1609   | 77.3        | 1897   | 6 CR610570  | CR610570 full-leng |
| 5          | 1609   | 77.3        | 2119   | 6 CR615336  | CR615336 full-leng |
| 6          | 1609   | 77.3        | 2124   | 6 CR593340  | CR593340 full-leng |
| 7          | 1594   | 76.6        | 1760   | 6 CR597295  | CR597295 full-leng |
| 8          | 1594   | 76.6        | 3536   | 6 CR749620  | CR749620 Homo sapi |
| 9          | 1556   | 74.8        | 1094   | 10 W18183   | W18183 IMAGE:20087 |
| 10         | 1446   | 69.5        | 841    | 9 CX795595  | CX795595 JGI_CAAJ1 |
| c 11       | 1415   | 68.0        | 818    | 9 DN530922  | DN530922 1398786 M |
| 12         | 1410.5 | 67.8        | 993    | 8 CO579074  | CO579074 ILLUMIGEN |
| 13         | 1403   | 67.4        | 869    | 10 DT309300 | DT309300 JGI_CAAJ1 |
| 14         | 1387   | 66.7        | 2190   | 6 AK164048  | AK164048 Mus muscu |
| 15         | 1379   | 66.3        | 890    | 3 BQ651422  | BQ651422 AGENCOURT |
| 16         | 1368   | 65.7        | 844    | 10 DV929624 | DV929624 LB03017.C |
| 17         | 1362.5 | 65.5        | 931    | 1 AL536089  | AL536089 AL536089  |
| c 18       | 1336   | 64.2        | 946    | 1 AL521664  | AL521664 AL521664  |
| 19         | 1334.5 | 64.1        | 930    | 3 BQ651351  | BQ651351 AGENCOURT |
| c 20       | 1333   | 64.1        | 935    | 4 CA976264  | CA976264 AGENCOURT |
| 21         | 1318.5 | 63.4        | 1094   | 2 BM478839  | BM478839 AGENCOURT |
| 22         | 1317   | 63.3        | 774    | 4 CB45245   | CB45245 726514 MA  |
| 23         | 1316   | 63.2        | 812    | 8 CV557169  | CV557169 UI-M-H20- |
| 24         | 1315   | 63.2        | 863    | 10 DT280981 | DT280981 JGI_CAAV9 |
| 25         | 1312   | 63.0        | 749    | 4 CB450430  | CB450430 705088 MA |
| 26         | 1308   | 62.9        | 886    | 3 BQ654247  | BQ654247 AGENCOURT |
| 27         | 1296   | 62.3        | 747    | 2 BI558388  | BI558388 603240449 |
| 28         | 1291   | 62.0        | 800    | 10 DT154092 | DT154092 JGI_ANN03 |
| 29         | 1268   | 60.9        | 889    | 3 BQ673346  | BQ673346 AGENCOURT |
| 30         | 1266   | 60.8        | 927    | 3 BQ669839  | BQ669839 AGENCOURT |
| 31         | 1263.5 | 60.7        | 1275   | 13 CL963186 | CL963186 OsIFCC008 |
| 32         | 1259   | 60.5        | 1745   | 6 AY106255  | AY106255 Zea mays  |
| 33         | 1257.5 | 60.4        | 1752   | 6 AY108486  | AY108486 Zea mays  |
| 34         | 1256   | 60.4        | 785    | 5 CK468065  | CK468065 939430 MA |
| 35         | 1255   | 60.3        | 781    | 10 DV907419 | DV907419 LB0282.CR |
| 36         | 1253.5 | 60.2        | 1197   | 13 CL957492 | CL957492 OsIFCC000 |
| 37         | 1247   | 59.9        | 1967   | 6 AY103545  | AY103545 Zea mays  |
| c 38       | 1245   | 59.8        | 785    | 5 CK465964  | CK465964 937126 MA |
| 39         | 1240.5 | 59.6        | 755    | 2 EG210727  | EG210727 RS130273  |
| 40         | 1233   | 59.3        | 1666   | 6 AY104068  | AY104068 Zea mays  |
| 41         | 1232   | 59.2        | 1409   | 6 CNSOABT   | BX815323 Arabidops |
| 42         | 1232   | 59.2        | 1461   | 6 CNSOABT   | BX817800 Arabidops |
| 43         | 1231   | 59.2        | 1711   | 6 DQ241848  | DQ241848 Solanum t |
| 44         | 1228   | 59.0        | 1593   | 6 DQ252508  | DQ252508 Solanum t |
| 45         | 1225   | 58.9        | 1638   | 6 CNS09YSF  | BX831446 Arabidops |

ALIGNMENTS

RESULT 1  
AK170965  
LOCUS  
DEFINITION Mus musculus NOD-derived Cd11c +ve dendritic cells cDNA, RIKEN  
2312 bp  
mrna  
linear  
HTC 21-SEP-2005

full-length enriched library, clone: F630213P21 product: glycogen  
synthase kinase 3 beta, full insert sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AK170965  
AK170965.1 GI:74221330  
HTC; CAP trapper  
Mus musculus (house mouse)

REFERENCE

Carninci, P. and Hayashizaki, Y.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.

1  
AUTHORS

High-efficiency full-length cDNA cloning

JOURNAL  
PUBMED

Meth. Enzymol. 303, 19-44 (1999)

2  
AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL  
PUBMED

Genome Res. 10 (10), 1617-1630 (2000)

3  
AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, K., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-Format  
sequencing pipeline with 384 multicapillary sequencer

JOURNAL  
PUBMED

Genome Res. 10 (11), 1757-1771 (2000)

4  
AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuura, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuenli, P., Lewis, S., Matsuo, Y., Nikaado, I., Pesole, G.,  
Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D.A., Kaniya, M., Lee, N.H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.  
and Hayashizaki, Y.

CONSORTM

FANTOM Consortium

JOURNAL  
PUBMED

Nature 409 (6821), 685-690 (2001)

5  
AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Ooato, N., Saito, R., Suzuki, H., Yamanaka, I.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V.,  
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,  
Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T.,  
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,  
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,  
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,  
Sutcliffe, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,  
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,  
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,  
Arakawa, T., Sato, K., Shiraki, T., Waki, K., Kawai, J., Iishi, Y.,  
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,  
Rogers, J., Birney, E. and Hayashizaki, Y.  
FANTOM Consortium  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420 (6915), 563-573 (2002)  
12466851  
6  
Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C.,  
Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R.,  
Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R.,  
Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E.,  
Ambesi-Impombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L.,  
Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M.,  
Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R.,  
Crome, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G.,  
di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G.,  
Fletcher, C.P., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M.,  
Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E.,  
Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N.,  
Hill, D., Hummel, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T.,  
Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H.,  
Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K.,  
Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J.,  
Liuni, S., McWilliam, S., Madan, B., Madera, M., Marchionni, L.,  
Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S.,  
Morris, K., Mortagi-Tabar, S., Mulder, N., Nakano, N., Nakachi, H.,  
Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O.,  
Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavoni, G.,  
Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z.,  
Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A.,  
Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C.A., Seno, S.,  
Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D.,  
Sincalir, B., Sperling, S., Stupka, E., Sugiyama, K., Sulcane, R.,  
Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S.,  
Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R.,  
Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A.,

Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T.,  
Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A.,  
Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M.,  
Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M.,  
Kato, T., Kawai, H., Kawagashira, N., Kawashima, T., Kojima, M.,  
Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M.,  
Plesky, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K.,  
Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.  
PANTOM Consortium  
The transcriptional landscape of the mammalian genome  
Science 309 (5740), 1559-1563 (2005)  
16141072  
7  
Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M.,  
Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H.,  
Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T.,  
Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L.,  
Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A.,  
Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and  
Wahlestedt, C.  
RIKEN Genome Exploration Research Group  
Antisense transcription in the mammalian transcriptome  
Science 309 (5740), 1564-1566 (2005)  
16141073  
8 (bases 1 to 2312)  
Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K.,  
Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S.,  
Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N.,  
Nishihori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D.,  
Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watahiki, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of  
Physiocal and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa, 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,  
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics  
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome  
Trust/MRC building Addenbrookes Hospital Cambridge) whose  
assistance we gratefully acknowledge.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/  
Location/Qualifiers  
1. 2312  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NOD"  
/db\_xref="FANTOM\_DB:F630213p21"  
/db\_xref="taxon:10090"  
/clone="F630213p21"  
/cell\_type="NOD-derived CD11c +ve dendritic cells"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"

973. 2235  
/note="unnamed protein product: glycogen synthase kinase 3  
beta (MCD|MGI:1861437 GB|NM\_019827, evidence: BLASTN, 99%,  
match=2294)  
putative"  
/codon\_start=1  
/protein\_id="BAB42146.1"  
/db\_xref="GI:74221331"  
/translation="MSGRRPTTTFSAESCKPVOQPSAFGSMKVRSDKSGKVTVA"TP  
GGQPCRPQEVSYDTDKVINGSGFVYQAKLDSGELVAVIKVLQDKRFKRELOIYHR  
KLDPCNIVRLRYFHSYSGSKDEVDVNLVDVYVETVYVARIYSAKQTLFVLYVKL  
YMYQLFRSLAIYHFGICHRDIPKQNLDPDPTAVLKLCDFGSAKQLRGSPNYSYIC  
SYRYRAPELIFGATDVTSSIDVMSAGCVLAELLGQPIFGDSGVQLVEIKVLGTP  
TREQIRENNPNYTFEKFPOIKAHPTKVPFRPTPEALCALCSLLEVTPTARLTPLEA  
CAHSFDELDPNVKLPNDRDTPALFNFTTQELSSNPPLATILIPHARIQAAASPPA

Alignment Scores:  
Pred. No.: 1.08e-210 Length: 2312  
Score: 2015.00 Matches: 386  
Percent Similarity: 98.0% Conservative: 0  
Best Local Similarity: 98.0% Mismatches: 2  
Query Match: 96.8% Indels: 6  
DB: 6 Gaps: 1

US-10-733-816-2 (1-394) x AK170965 (1-2312)

QY 7 GLUGlyGlyGly-----MetSerGlyArgProArgThrThrSerPhe 20  
DB 943 GAAGGAAAGGCTGATTCAGGAAGAGCCATCATCTGCGGGCGCAGCAACCTCTTT 1002  
QY 21 AlaGluSerCysAlaProValGlnGlnProSerAlaAlaPheClySerMetLysValSerArg 40  
DB 1003 GCGGAGAGCTGCAAGCCAGTCAGCAGCCCTTTCAGCTTTTGTAGTCATCAAGTTACAGA 1062  
QY 41 AspLysAspGlySerLysValThrValValAlaThrProGlyGlnGlyProAspArg 60  
DB 1063 GATAAAGATGCGACGACGATTAACCACTGATGCGCACTCTCTGGCAGGGTCTTGACAGG 1122  
QY 61 ProGlnGluValSerThrAspThrLysValLleGlyAsnGlySerPheGlyValVal 80  
DB 1123 CCACAGGAAGTCAGTTATACAGACACGAAAGTGATTGGAATGGATCATTTGGTGTGTA 1182  
QY 81 TyrGlnAlaLysLeuCysAspSerGlyGluLeuValAlaLleLysLysValLeuGlnAsp 100  
DB 1183 TATCAAGCCAACTTTGTGATTCTCGGAGAACTGGTTGCCATCAAGAAAGTTCTACAGGAC 1242  
QY 101 LysArgPheLysAsnArgGluLeuGlnLleMetArgLysLeuAspHisCysAsnLleVal 120  
DB 1243 AAGCGATTAAAGAACCGAGCTCCAGATCATGAGAAAGCTAGACCCTGTAAACATAGTC 1302  
QY 121 AsgLewArgTyrPhePheThrSerSerGlyGluLysLysAspGluValTyrLeuAsnLeu 140  
DB 1303 CGACTGCGGTATTTCTTCTACTGAGTGGCGAGAGAAAGNTGAGGTCTACTTACCTG 1362  
QY 141 ValLeuAspTyrValProGluThrValTyrArgValAlaAlaArgHisTyrSerArgAlaLys 160  
DB 1363 GTGCTGGACTATGTTCCGAGACAGAGTGTACAGAGTCCGACAGACTATAGTCCAGCAAG 1422  
QY 161 GlnThrLeuProValLleTyrValLysLeuTyrMetTyrGlnLeuPheArgSerLeuAla 180  
DB 1423 CAGACACTCCCTGATCTATGTCAAGTTGTATATGTATGATCATGCTTTCAGAAAGTGTAGCC 1482

FEATURES  
source

COMMENT

TITLE  
JOURNAL

REFERENCE  
AUTHORS

CONSRM  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS

|            |      |   |      |
|------------|------|---|------|
| QY         | 181  | TyrileHisSerPheGlyIleCysHisArgAspIleLysProGlnAsnLeuLeuAsp         | 200  |
| DB         | 1483 | TATATCAATCTCTTGGAAATCTGCATCGACATTAACACACACACCTCTTGTTGGAT          | 1542 |
| QY         | 201  | ProAspThrAlaValLeuLysLeuCysAspPheGlySerAlaLysGlnLeuValArgGly      | 220  |
| DB         | 1543 | CTGTATACAGCTGTATTAAACTCTGACTTTGGAAGTGCAAGACAGCTGTGTCGAGGA         | 1602 |
| QY         | 221  | GluProAsnValSerTyrIleCysSerArgTyrTyrArgAlaProGluLeuIlePheGly      | 240  |
| DB         | 1603 | GAGCCCAATCTTATATATCTCTCGTACTACAGGACACAGAGTGTGATCTTTGGA            | 1662 |
| QY         | 241  | AlaThrAspTyrThrSerIleAspValTrpSerAlaGlyCysValLeuAlaGluLeu         | 260  |
| DB         | 1663 | GCCACTGATTACAGCTCCAGTATAGATGTGTCTCGAGCTGTGTGTGGCTGAATG            | 1722 |
| QY         | 261  | LeuLeuGlyGlnProIlePheProGlyAspSerGlyValAspGlnLeuValGluIleIle      | 280  |
| DB         | 1723 | TGTCTAGACACCAATATTTCTTGGGACAGTGTGTGTGATCAGTGTGTGGAAATATA          | 1782 |
| QY         | 281  | LysValLeuGlyThrProThrArgGluGlnIleArgGluMetAsnProAsnTyrThrGlu      | 300  |
| DB         | 1783 | RAGTCTAGAACACCAACAGGAGCAATATAGAAATAGAACCAATATATACAGAA             | 1842 |
| QY         | 301  | PheLysPheProGlnIleLysAlaHisProTrpThrLysValPheArgProArgThrPro      | 320  |
| DB         | 1843 | TTCAAAATCCCTCAATCAAGGCACATCTTGGCAAGAGTCTTCCGGCCCCCAACTCA          | 1902 |
| QY         | 321  | ProGluAlaIleAlaLeuLysSerArgLeuLeuGluTyrThrProThrAlaArgLeuThr      | 340  |
| DB         | 1903 | CCAGAGCAATATGCACTGTGCAGCCCTCTGTCTGGAGTACACACCTACCGCCGCTACA        | 1962 |
| QY         | 341  | ProLeuGluAlaCysAlaHisSerPhePheAspGluLeuArgAspProAsnValLysHis      | 360  |
| DB         | 1963 | CCACTGGAACTTGTGCACATTCATTTTCGATGAATTCGGGACCCCAATGTCAAACTA         | 2022 |
| QY         | 361  | ProAsnGlyArgAspThrProAlaLeuPheAsnPheThrThrGlnGluLeuSerSerAsn      | 380  |
| DB         | 2023 | CCAAATGGCGAGACACACCTGCACTCTTCAACTTTTACCACCTCAAGAACTGTCAAGTAA      | 2082 |
| QY         | 381  | ProProLeuAlaThrIleLeuIleProProHisAlaArgIle                        | 394  |
| DB         | 2083 | CCCCCTGTGCCACCATCTTATCCCTCCACATCTCGGAT                            | 2124 |
| RESULT 2   |      |   |      |
| AK154293   |      | 2847 bp mRNA linear HTC 21-SEP-2005                               |      |
| LOCUS      |      | Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, RIKEN    |      |
| DEFINITION |      | full-length enriched library, clone:F630015J08 product:glycogen   |      |
|            |      | synthase kinase 3 beta, full insert sequence.                     |      |
| ACCESSION  |      | AK154293  |      |
| VERSION    |      | AK154293.1 GI:74178475  |      |
| KEYWORDS   |      | HTC; CAP trapper.   |      |
| SOURCE     |      | Mus musculus (house mouse)  |      |
| ORGANISM   |      | Mus musculus  |      |
|            |      | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |      |
|            |      | Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;           |      |
|            |      | Sciurognathi; Muridae; Murinae; Mus.                              |      |
| REFERENCE  |      | 1   |      |
| AUTHORS    |      | Carninci, P. and Hayashizaki, Y.                                  |      |

High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

11076861

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, F., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.

RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

11217851

5

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusci, V., Choithia, C., Corbani, L. E., Cousins, S., Della, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

Numata,K., Okido,T., Pavan,W.J., Perteau,G., Pesole,G.,  
 Petrovsky,N., Pillai,R., Pontius,J.U., Oi,D., Ramachandran,S.,  
 Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,  
 Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,  
 Saitana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,  
 Vekurova,K., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,  
 Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,  
 Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,  
 Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,  
 Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,  
 Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,  
 Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,K., Shibata,K.,  
 Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,  
 Rogers,J., Birney,E. and Hayashizaki,Y.  
 PANTOM Consortium  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420 (6915), 563-573 (2002)  
 12466851  
 6  
 Carninci,P., Kasukawa,T., Katayama,S., Gough,J., Frith,M.C.,  
 Maeda,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C., Kodzius,R.,  
 Shimokawa,K., Bajic,V.B., Brenner,S.E., Batalov,S., Forrest,A.R.,  
 Zavolan,M., Davis,M.J., Wilming,L.G., Aidinis,V., Allen,J.E.,  
 Ambesi-Impombato,A., Apweiler,R., Aturaliya,R.N., Bailey,T.L.,  
 Bansal,M., Baxter,L., Beisel,K.W., Bersano,T., Bono,H., Chalk,A.M.,  
 Chiu,K.P., Choudhary,V., Christoffels,A., Clutterbuck,D.R.,  
 Crome,M.E., Dalla,E., Dalrymple,B.P., de Bono,B., Della Gatta,G.,  
 di Bernardo,D., Down,T., Engstrom,P., Fagioli,M., Faulkner,G.,  
 Fletcher,C.F., Fukushima,T., Furuno,M., Futaki,S., Gariboldi,M.,  
 Georgii-Hemming,P., Gingeras,T.R., Gojobori,T., Green,R.E.,  
 Guinichich,S., Harbers,M., Hayashi,Y., Hensch,T.K., Hirokawa,N.,  
 Hill,D., Humincik,L., Iacono,M., Ikeo,K., Iwama,A., Ishikawa,T.,  
 Jakt,M., Kanapin,A., Katoh,M., Kawasawa,Y., Kello,J., Kitamura,H.,  
 Kitano,H., Kollas,G., Kriehnan,S.P., Kruger,A., Kummerfeld,S.K.,  
 Kurochkin,I.V., Larau,L.F., Lazarevic,D., Lipovich,L., Liu,J.,  
 Liuni,S., McWilliam,S., Madan Babu,M., Madera,M., Marchionni,L.,  
 Matsuda,H., Matsuzawa,S., Miki,H., Mignone,F., Miyake,S.,  
 Morris,K., Mottagui-Tabar,S., Mulder,N., Nakano,N., Nakachi,H.,  
 Ng,P., Nilsson,R., Nishiguchi,S., Nishikawa,S., Nori,F., Ohara,O.,  
 Okazaki,Y., Orlando,V., Pang,K.C., Pavan,W.J., Pavesi,G.,  
 Pesole,G., Petrovsky,N., Piazza,S., Reed,J., Reid,J.F., Ring,B.Z.,  
 Ringwald,M., Rost,B., Ruan,Y., Salzberg,S.L., Sandelin,A.,  
 Schneider,C., Schonbach,C., Sekiguchi,K., Semple,C.A., Seno,S.,  
 Sessa,L., Sheng,Y., Shibata,Y., Shimada,H., Shimada,K., Silva,D.,  
 Sinclair,B., Sperling,S., Stupka,E., Sugura,K., Sultana,R.,  
 Takenaka,Y., Taki,K., Tamoja,K., Tan,S.L., Tang,S., Taylor,M.S.,  
 Tegner,J., Teichmann,S.A., Ueda,H.R., van Nimwegen,B., Verardo,R.,  
 Wei,C.L., Yagi,K., Yamanishi,H., Zabarovsky,E., Zhu,S., Zimmer,A.,  
 Hide,W., Bult,C., Grimmond,S.M., Teasdale,R.D., Liu,E.T.,  
 Brusic,V., Quackenbush,J., Wahlestedt,C., Mattick,J.S., Hume,D.A.,  
 Kai,C., Sasaki,D., Tomaru,Y., Fukuda,S., Kanamori-Katayama,M.,  
 Suzuki,M., Aoki,J., Arakawa,T., Iida,J., Imamura,K., Itoh,M.,  
 Kato,T., Kawaji,H., Kawagashira,N., Kawashima,T., Kojima,M.,  
 Kondo,S., Konno,H., Nakano,K., Ninomiya,N., Nishio,T., Okada,M.,  
 Plessky,C., Shibata,K., Shiraki,T., Suzuki,T., Tagami,M., Waki,K.,  
 Watanishi,A., Okamura-Oho,Y., Suzuki,H., Kawai,J. and Hayashizaki,Y.  
 PANTOM Consortium  
 The transcriptional landscape of the mammalian genome  
 Science 309 (5740), 1559-1563 (2005)  
 16141072  
 7  
 Katayama,S., Tomaru,Y., Kasukawa,T., Waki,K., Nakanishi,M.,  
 Nakamura,M., Nishida,H., Yap,C.C., Suzuki,M., Kawai,J., Suzuki,H.,  
 Carninci,P., Hayashizaki,Y., Wells,C., Frith,M., Ravasi,T.,  
 Pang,K.C., Hallinan,J., Mattick,J., Hume,D.A., Lipovich,L.,  
 Batalov,S., Engstrom,P.G., Mizuno,Y., Faghihi,M.A., Sandelin,A.,  
 Chalk,A.M., Mottagui-Tabar,S., Liang,Z., Lenhard,B. and  
 Wahlestedt,C.  
 RIKEN Genome Exploration Research Group  
 Antisense transcription in the mammalian transcriptome  
 Science 309 (5740), 1564-1566 (2005)  
 16141073  
 8 (bases 1 to 2847)  
 Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K.,  
 Hori,F., Iida,J., Imamura,K., Imotani,K., Itoh,M., Kanagawa,S.,  
 Kawai,J., Kojima,M., Konno,H., Murata,M., Nakamura,M., Ninomiya,N.,  
 Nishiyori,H., Nomura,K., Ohno,M., Sakazume,N., Sano,H., Sasaki,D.,  
 Shibata,K., Shiraki,T., Tagami,M., Tagami,Y., Waki,K., Watanishi,A.,  
 Muramatsu,M. and Hayashizaki,Y.  
 Direct Submission  
 Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa, 230-0045, Japan (E-mail:genome-res@gsr.riken.jp,  
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,  
 Fax:81-45-503-9216)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in Riken.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics  
 Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome  
 Trust/MRC building Addenbrookes Hospital Cambridge) whose  
 assistance we gratefully acknowledge.  
 Please visit our web site for further details.  
 URL:http://genome.gsc.riken.jp/  
 URL:http://fantom.gsc.riken.jp/  
 Location/Qualifiers  
 1. 2847  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="NOD"  
 /db\_xref="FANTOM DB:F630015J08"  
 /db\_xref="taxon:10090"  
 /clone="F630015J08"  
 /cell\_type="NOD-derived CD11c +ve dendritic cells"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 1004. 2266  
 /note="unnamed protein product; glycogen synthase kinase 3  
 beta (MGD|MG1:1861437 GB|NM\_019827, evidence: BLASTN, 99%,  
 match=2350)  
 putative"  
 /codon\_start=1  
 /protein\_id="BA832494.1"  
 /db\_xref="GI:74178476"  
 /translation="MSGRPTTSFAESCKPVQPSAFGSMKVSBDKSGKVTTVVAP  
 GQGPDRPQEVSTDTKIVNGSFGVYQAKLDSGLVAIKKVLQDKRFRKREIQVWR  
 KLDHCNIVLFFYFSGSGKDEVYLNLDVYVPETVYRVARHYSRKAQTLPVYVKL  
 YMYQLFRSLAYHSFGICHDRDKPQNLDDPTAVLKLCDFGSAKQLVRGEPNVSXIC

SRYYRAPELIFGATDYTSIDWSAGCVLAELLGQIFPDGSDVQLVEIHKVLGTP  
 TREOIREKNPNYTEPKFQIKAHPTKVRPPTPEALALCSLLEYTPTRALTPLEA  
 CAHSFDELDPNWKLPNGRDTPLAFNFTTQELSSNPPLATILIPPHARIOAAAPPA

## Alignment Scores:

Pred. No.: 1.47e-210 Length: 2847  
 Score: 2015.00 Matches: 386  
 Percent Similarity: 98.0% Conservativity: 0  
 Best Local Similarity: 98.0% Mismatches: 2  
 Query Match: 96.8% Indels: 6  
 DB: Gaps: 1

US-10-733-816-2 (1-394) x AK154293 (1-2847)

QY 7 GluGlyGlyCly-----MetSerClyArgProArgThrThrSerPhe 20  
 ||||| |||  
 DB 974 GAAGGAAAGGTGATTCAAGAGAGCCATCATGTGCGGGCGACCGAGAACCACTCTCTT 1033  
 QY 21 AlaGluSerCysAlaProValGlnGlnProSerAlaPheGlySerMetLysValSerArg 40  
 ||||| |||||  
 DB 1034 CGCGAGAGCTGCAGCCAGTCGACAGCCCTTCAGCTTTTGGTAGCATGAAGTTACAGGA 1093  
 QY 41 AspLysAspGlySerLysValThrValValAlaThrProGlyGlnGlyProAspArg 60  
 ||||| |||||  
 DB 1094 GATTAAGATGGCCAGGTAACACAGTAGTGCAACTCTCTGCCAGGGTCTTCACAGG 1153  
 QY 61 ProGlnGluValSerTyrThrAspThrLysValIleGlyAsnGlySerPheGlyValVal 80  
 ||||| |||||  
 DB 1154 CCACAGGAGTCAGTTATACAGACACGAAAGTGATTGGAAATGGATCAATTTGGTGTGTA 1213  
 QY 81 TyrGlnAlaLysCysAspSerGlyGluLeuValAlaIleLysValLysValLeuGlnAsp 100  
 ||||| |||||  
 DB 1214 TATCAAGCCAAATCTTGTGATCTCTGGAGAACTGTTGCCATCAAGAAAGTTCTACAGGAC 1273  
 QY 101 LysArgPheLysAsnArgGluLeuGlnIleMetArgLysLeuAspHisCysAsnIleVal 120  
 ||||| |||||  
 DB 1274 AAGCGATTAAAGAACCGAGAGCTCCAGATCATGAGAAAGCTAGACCACTGTACATAGTC 1333  
 QY 121 ArgLeuArgTyrPhePheTyrSerSerGlyGluLysLysAspGluValTyrLeuAsnLeu 140  
 ||||| |||||  
 DB 1334 CGACTGCGGTATTTCTTACTACTGAGTGGCGAGGAAGATGAGTCTTACCTTAACCTG 1393  
 QY 141 ValLeuAspTyrValProGluThrValTyrArgValAlaArgHisTyrSerArgAlaLys 160  
 ||||| |||||  
 DB 1394 GTGTGACTATGTTCCGAGACAGTGTACAGAGTCCAGACACTATAGTCGAGGCAAG 1453  
 QY 161 GlnThrLeuProValIleTyrValLysLeuTyrMetTyrGlnLeuPheArgSerLeuAla 180  
 ||||| |||||  
 DB 1454 CAGACACTCCCTGTGATCTATGTCAAAGTTGTATATGTATCATGCTTTTCAGAAAGCTTAGCC 1513  
 QY 181 TyrIleHisSerPheGlyIleCysHisArgAspIleLysProGlnAsnLeuLeuAsp 200  
 ||||| |||||  
 DB 1514 TATATCAATCTCTTGGAAATCTGCCATCGACATTAACCCACAGAACCTCTTGTGTGAT 1573  
 QY 201 ProAspThrAlaValLeuLysLeuCysAspPheGlySerAlaLysGlnLeuValArgGly 220  
 ||||| |||||  
 DB 1574 CCTGATACAGTGTATTAACACTCTGACTTTGGAAGTGCAAGACAGCTGGTCCGAGGA 1633  
 QY 221 GluProAsnValSerTyrIleCysSerArgTyrTyrArgAlaProGluLeuIlePheGly 240  
 ||||| |||||  
 DB 1634 GAGGCCAAATGTTTCATATATCTGTCTCGTACTACAGGGCCACAGAGTGTATCTTTGGA 1693

QY 241 AlaThrAspTyrThrSerSerIleAspValTrpSerAlaGlyCysValLeuAlaGluLeu 260  
 ||||| |||||  
 DB 1694 GCCACTGATTACACGTCCAGTATGATGTATGGTCTGCAGAGCTGTGTGTGGCTGAATTG 1753  
 QY 261 LeuLeuGlyGlnProIlePhePheProGlyAspSerGlyValAspGlnLeuValGluIle 280  
 ||||| |||||  
 DB 1754 TTGCTAGGACACCAATATTTCTGGGACAGTGGTGTGGATCAGTGTGGGAATATA 1813  
 QY 281 LysValLeuGlyThrProThrArgGluGlnIleArgGluMetAsnProAsnTyrThrGlu 300  
 ||||| |||||  
 DB 1814 AAGGTCCTAGGAACACCAACAGGAGCAAAATAGAGAAATGAACCAAAATTATACAGA 1873  
 QY 301 PheLysPheProGlnIleLysAlaHisProTyrThrLysValPheArgProAspThrPro 320  
 ||||| |||||  
 DB 1874 TTCAAAATTCCTCAAAATCAAGGCACATCTCTTGGACAAAGTCTTCCGGCCCCCAACTCA 1933  
 QY 321 ProGluAlaIleAlaLeuCysSerArgLeuLeuGluTyrThrProThrAlaArgLeuThr 340  
 ||||| |||||  
 DB 1934 CCAGAGCAATTGCACTGTGCGCCGCTCTGTGGGTACACACTTCCGCCCGGCTTACA 1993  
 QY 341 ProLeuGluAlaCysAlaHisSerPhePheAspGluLeuArgAspProAsnValLysHis 360  
 ||||| |||||  
 DB 1994 CCACCTGGAAGCTTGTGCACATTCATTTTTCGATGAATGGGAGCCCAATGTCAAACTA 2053  
 QY 361 ProAsnGlyArgAspThrProAlaLeuPheAsnPheThrThrGlnGluLeuSerSerAsn 380  
 ||||| |||||  
 DB 2054 CCAAATGGCGAGACACACCTTGCATCTTCAACTTTTACCACCTCAAGAACTGTCAAGTAAC 2113  
 QY 381 ProProLeuAlaThrIleLeuIleProProHisAlaArgIle 394  
 ||||| |||||  
 DB 2114 CCCCCTGTGCCACCATCTTATCCCTCCACATCTCGGATT 2155

## RESULT 3

LOCUS CR617019 1732 bp mRNA linear HTC 21-JUL-2004  
 DEFINITION full-length cDNA clone CS0DB003YJ02 of Neuroblastoma Cot  
 10-normalized of Homo sapiens (human).

ACCESSION CR617019  
 VERSION CR617019.1 GI:50497826  
 KEYWORDS HTC; CNSLT\_cDNA.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

REFERENCE 1 (bases 1 to 1732)  
 AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue

REFERENCE 2 (bases 1 to 1732)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT - Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

# FEATURES

## Location/Qualifiers

1..1732  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS008003YJ02"  
 /tissue\_type="Neuroblastoma Cot 10-normalized"  
 /plasmid="pCMVSPORT\_6"

## ORIGIN

### Alignment Scores:

Pred. No.: 1.18e-210 Length: 1732  
 Score: 2013.00 Matches: 383  
 Percent Similarity: 99.7% Conservative: 0  
 Best Local Similarity: 99.7% Mismatches: 1  
 Query Match: 96.7% Indels: 0  
 DB: 6 Gaps: 0

US-10-733-816-2 (1-394) x CR610570 (1-1732)

QY 11 MetSerGlyArgProArgThrThrSerPheAlaGluSerCysLysProValGlnGlnPro 30  
 DB 395 ATGTCAGGCGCGCCAGAACCACTCTCTTTGGGAGAGCTGCAACCGGTGCGAGAGCT 454  
 QY 31 SerAlaPheGlySerMetLysValSerArgAspLysAspGlySerLysValThrVal 50  
 DB 455 TCAGCTTTGGCAGCATCAAGTTACGAGACAAAGCGGCGGCGGAGGTGACACAGTG 514  
 QY 51 ValAlaThrProGlnGlnProAspArgProGlnGlnValSerThrAspThrLys 70  
 DB 515 GTGGCAACTCTCTGGGCGGCTCCAGACAGGCGCACAGAGTCAGTCATACAGACACTAA 574  
 QY 71 ValIleGlyAsnGlySerPheGlyValValTyrGlnAlaLysLeuCysAspSerGlyGlu 90  
 DB 575 GTGATTGGAAATGGATCATTTGGTGTGTATATCAAGCCCAAACTTTGTGATTGAGGAA 634  
 QY 91 LeuValAlaIleLysValLeuGlnAspLysArgPheLysAsnArgGluLeuGlnIle 110  
 DB 635 CTGGTCCCATCAAGAAAGTATTGGCAGACAGAGATTAGAAATCGAGAGCTCCAGATC 694  
 QY 111 MetArgLysLeuAspHisCysAsnIleValArgLeuArgTyrPhePheTyrSerSerGly 130  
 DB 695 ATGAGAAAGCTAGATCACTGTAAATAGTCCGATTGGATTGCTTCTTCTACTCCAGTGT 754  
 QY 131 GluLysLysAspGluValTyrLeuAsnLeuValLeuAspTyrValProGluThrValTyr 150  
 DB 755 GAGAAGAAAGATGAGTCTATCTTAATCTGGTGTGCTGACTATGTTCCGAAACAGATATAC 814  
 QY 151 ArgValAlaAlaArgHisTyrSerArgAlaLysGlnThrLeuProValIleTyrValLysLeu 170  
 DB 815 AGAGTTCCAGACACTAGTCGAGCGCAACAGACGCTCCCTGTGATTATGTCAGATGTG 874  
 QY 171 TyrMetTyrGlnLeuPheArgSerLeuAlaTyrIleHisSerPheGlyIleCysHisArg 190  
 DB 875 TATATGATGCTGTCTCGAAGTTTAGCTTAATCATCTCTTTGGAACTGCCATCG 934  
 QY 191 AspIleLysProGlnAsnLeuLeuAspProAspThrAlaValLeuLysLeuCysAsp 210

DB 935 GATATTAACCCGACAGACCTCTTTGGATCTCTGATCTACTGCTGTATTAAACCTCTGTGAC 994  
 QY 211 PheGlySerAlaLysGlnLeuValArgGlyGluProAsnValSerTyrIleCysSerArg 230  
 DB 995 TTTGGAAGTGCAGACAGCTGTCCGAGGAGAACCCCAATGTTTCGTATATCTGTCTCGG 1054  
 QY 231 TyrTyrArgAlaProGluLeuIlePheGlyAlaThrAspTyrThrSerSerIleAspVal 250  
 DB 1055 TACTATAGGCGACAGAGTGTATCTTTGGAGCCACTGATTATACCTCTAGTAGATGTA 1114  
 QY 251 TrpSerAlaGlyCysValLeuAlaGluLeuLeuGlyGlnProIlePheProGlyAsp 270  
 DB 1115 TGGTCTGCTGCTGTGTGTGGCTGAGCTGTTACTAGGCAACCAATATTTCCAGGGGAT 1174  
 QY 271 SerGlyValAspGlnLeuValGluIleLysValLeuGlyThrProThrArgGluGln 290  
 DB 1175 AGTGTGTGGATCAGTTGGTAGAAATAATCAAGGTCTCTGGAACTCCCAACAGGGAGCA 1234  
 QY 291 IleArgGluMetAsnProAsnTyrThrGluPheLysPheProGlnIleLysAlaHisPro 310  
 DB 1235 ATCAGAGAAATGAACCCAACTACACAGAAATTTAAATTCCTCAATATTAGGCACATCT 1294  
 QY 311 TrpThrLysValPheArgProArgThrProProGluAlaIleAlaLeuCysSerArgLeu 330  
 DB 1295 TGGACTAAGTCTTCGACCCCGAACTCCACGCGAGGCAATTGCACTGTGTAGCCCTCTG 1354  
 QY 331 LeuGluTyrThrProThrAlaArgLeuThrProLeuGluAlaCysAlaHisSerPhePhe 350  
 DB 1355 CTGGAGTATACCAACTGCCCGCACTAACCACTTGGAGCTTGTGCACATTCATTTT 1414  
 QY 351 AspGluLeuArgAspProAsnValLysHisProAsnGlyArgAspThrProAlaLeuPhe 370  
 DB 1415 GATGAATTTACGGGACCCAAATGTCAAACTACCAATGGCGAGACACACCTGTGACTTTC 1474  
 QY 371 AsnPheThrGlnGluLeuSerSerAsnProProLeuAlaThrIleLeuLeuProPro 390  
 DB 1475 AACTTCACCACTCAAGAACTGTCAAGTAATCCACTCTGCTGCTACCATCTTATTCCTCT 1534  
 QY 391 HisAlaArgIle 394  
 DB 1535 CATGCTCGGATT 1546  
 RESULT 4  
 CR610570  
 LOCUS 1897 bp mRNA linear HTC 21-JUL-2004  
 full-length cDNA clone CS0DJ012YC04 of T cells (Jurkat cell line)  
 DEFINITION Cot 10-normalized of Homo sapiens (human).  
 ACCESSION CR610570  
 VERSION CR610570.1 GI:50491377  
 KEYWORDS HTC; CNSLT\_cDNA.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1897)  
 AUTHORS Li W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 CONTACT : Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

REFERENCE TITLE JOURNAL COMMENT FEATURES ORIGIN

Paraday Avenue  
2 (bases 1 to 1897)  
Genoscope.  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191, 91006 EVRY cedex - FRANCE (E-mail : secre@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.

Location/Qualifiers  
1. 1897  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DJ012YC04"  
/tissue\_type="T cells (Jurkat cell line) Cot  
10-normalized"  
/plasmid="pCMVSPORT\_6"

Alignment Scores:  
Pred. No.: 5,52e-166 Length: 1897  
Score: 1609.00 Matches: 319  
Percent Similarity: 77.3% Conservative: 24  
Best Local Similarity: 71.8% Mismatches: 43  
Query Match: 77.3% Indels: 58  
DB: 6 Gaps: 4

US-10-733-816-2 (1-394) x CR610570 (1-1897)

QY 5 ProMetGluGlyGly---GlyMetSerGlyArgProArgThrThrSerPheAlaGluSer 23  
DB 176 CCTTCGGAGGCGGCCCTCGGGGCTCGGCGAGGCGCGGACTAGCTGTTCCGGAG--- 232  
QY 24 CysLysPro-----ValGlnGln 29  
DB 233 -----CCCGCGCGGAGGCGGAGGCGGCGCGCGCGCGGAGGCTCGGCTCCGGC 286  
QY 30 ProSerAlaPheGlySerMetLys----- 37  
DB 287 CCAGCGCGCACCGCGCGGGAAGGCATCTGTGCGGCGCATGSGTGGGCGCTCGGCGCC 346  
QY 37 ----- 37  
DB 347 TCAGAGCTTCGGGGTGGACCCGCGCGCACCGCGCGGAGGCGGAGCGCGCGCGCA 406  
QY 38 -----ValSerArgAspLysAspGlySerLysValThrThr 49  
DB 407 GGCAGTAGCTTCCCGCGCGCGGGGTGAAGCTGGCGCGTGCACGCGGAGGTTGACCA 466  
QY 50 ValValAlaThrProGlyGlnGlyProAspArgProGlnGluValSerThrAspThr 69  
DB 467 GTCTAGCCACTCTAGGCCAAGGCCAGAGCGCTCCCAAGAGTGGCTTACACGACATC 526  
QY 70 LysValIleGlyAsnGlySerPheGlyValValThrGlnAlaLysLysCysAspSerGly 89  
DB 527 AAGGTGATTGGCAATGGGCTCATTTTGGGGTGGTGTACCGAGCGCGCTCGGACGACCGAG 586

QY 90 GluLeuValAlaIleLysValLysValLeuGlnAspLysArgPheLysAsnArgGluLeuGln 109  
DB 587 GAAGTAGTCCCATCAGAGGTTCTCCAGGACGAGGTTTCAGACCCAGAGCTGCAG 546  
QY 110 IleMetArgLysLeuAspHisCysAsnIleValArgLeuArgTyrPhePheTyrSer 129  
DB 647 ATCATGCTGAAGCTGCACCACTATTTGGAGGCTGAGATACTTTTCTACTCCAGT 706  
QY 130 GlyGluLysLysAspGluValTyrLeuAsnLeuValLeuAspTyrValProGluThrVal 149  
DB 707 GCGGAGAAGAACGACGCTTTACCTAAATCTGGTCTGGAATATGTGCCGAGACAGTG 766  
QY 150 TyrArgValAlaArgHisTyrSerArgAlaLysGlnThrLeuProValIleTyrValLys 169  
DB 767 TACCGGTGGCGCCCACTTCACCAAGGCCAAGTTGACCATCCCTTCTCTATGTCAG 826  
QY 170 LeuTyrMetTyrGlnLeuPheArgSerLeuAlaTyrIleHisSerPheGlyIleCysHis 189  
DB 827 GTGTACATGTACCAGCTCTTCGACAGCTTGGCTTACATCCACTCCAGGCGGTGTGCAC 886  
QY 190 ArgAspIleLysProGlnAsnLeuLeuLeuAspProAspThrAlaValLeuLysLeuCys 209  
DB 887 CGCGACATCAAGCCCGCAGACCTGCTGGTGGACCTGCACACTGCTGCTCAAGCTCTGC 946  
QY 210 AspPheGlySerAlaLysGlnLeuValArgGlyGluProAsnValSerTyrIleCysSer 229  
DB 947 GATTTGGCAGTGCACCAAGCACTGGTCCGAGGGGAGGCCCAATGTCTCTACATCTCTCT 1006  
QY 230 ArgTyrTyrArgAlaProGluLeuIlePheGlyAlaThrAspTyrThrSerSerIleAsp 249  
DB 1007 CCTACTACCGGCGCCAGAGCTCATTTTGGAGCCACTGATTACACCTCATCATCATCAT 1066  
QY 250 ValTrpSerAlaGlyCysValLeuAlaGluLeuLeuGlyGlnProIlePheProGly 269  
DB 1067 GTTGTGCTGCTGCTGTACTGGCAGAGCTCTCTTGGCCAGCCCATCTTCCCTGGG 1146  
QY 270 AspSerGlyValAspGlnLeuValGluIleLysValLeuGlyThrProThrArgGlu 289  
DB 1127 GACAGTGGGTGGACCACTGCTGGAGATCATCAAGGTGCTGGGAACCAACCCCGGAA 1186  
QY 290 GlnIleArgGluMetAsnProAsnTyrThrGluPheLysPheProGlnIleLysAlaHis 309  
DB 1187 CAATTCGAGAGATGAACCCCACTACGAGGTTCAAGTTCCCTCAGATTAAAGCTCAC 1246  
QY 310 ProTyrThrLysValPheArgProArgThrProGluAlaIleAlaLeuCysSerArg 329  
DB 1247 CCTGGCAAGAGTGTTCAAATCTCGNACGCCGCGAGGCGCATCGGCTCTGCTCTAGC 1306  
QY 330 LeuLeuGluTyrThrProThrAlaArgLeuThrProLeuGluAlaCysAlaHisSerPhe 349  
DB 1307 CTGCTGGAGTACACCCCATCTCAAGGCTCTCCCACTAGAGGCTGTGGCAGACGCTTC 1346  
QY 350 PheAspGluLeuArgAspProAsnValLysHisProAsnGlyArgAspThrProAlaLeu 369  
DB 1367 TTTGATCAACTGGCATGTCTCGGAACCCAGCTGCTTAACAACCGCCCATCTTCCCTCTC 1426  
QY 370 PheAsnPheThrThrGlnGluLeuSerSerAsnProProLeuAlaThrIleLeuLeuPro 389  
DB 1427 TTCAACTTCAGTGTGGTGAATCTTCATCCAACCGTCTCTCAACGCCATTTTATTCCT 1486  
QY 390 ProHisAlaArg 393



CR615336

## DEFINITION

ACCESSION

## KEYWORDS

ORGANIZATIONAL

## REFERENCES

**ATTENTION**

REMARK:

RECEIVED  
MAY 19 1964

77111

**COMMENT**

**000000**

**BON**

•

1111

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

Pred. No

Percent

## Query Ma

|    |  |      |  |
|----|--|------|--|
| QY |  | 30   | ProSerAlaPheGlySerMetLys-----37                                      |
| DB |  | 219  | CCAGCGCGCACCCGGCGCGGAAGAAGCATCTGTCTCGGGGCCATGGGTGGGGGCGTGTGGGGGC 278 |
| QY |  | 37   | -----37  |
| DB |  | 279  | TGAGACTCCGGGGTGACCCGGCGCAGCGCGGAGGACGACGGAGGCCGCCGCGCGCA 33          |
| QY |  | 38   | -----ValSerArgAspLysAspGlySerLysValThr 49                            |
| DB |  | 339  | GSCACTAGCTTCGCCGCCGCCGGGTGAAGCTGGCGCGCTGACAGCGGGAAGTGATCCACA 39      |
| QY |  | 50   | ValValAlaThrProGlyGlnGlyProAspArgProGlnGluValSerTyThrAspThr 69       |
| DB |  | 399  | GTCTAGCCACTCTTAGCGCCAAAGCCAGACGGCTCCCAAGAAAGTGGCTTTACACGGACATC 45    |
| QY |  | 70   | LysValIleGlyAsnGlySerPheGlyValValTyGlnAlaLysLeuCysAspSerGly 69       |
| DB |  | 459  | AAGTGNATTGGCATGGCTCAHTTGGGTGTGTACCAGGCGCTGGCAGAGACCCNGG 51           |
| QY |  | 90   | GluLeuValAlaIleLysLysValLeuGlnAspLysArgPheLysAsnArgGluLeuGln 10      |
| DB |  | 519  | GNACTAGTCGCATCAGAGAGTTCTCCAGGACAGAGGTTCAAGAACCGAGAGCTCCAG 57         |
| QY |  | 110  | IleMetArgLysLeuAspHisCysAsnIleValArgLeuArgTyrPhePheTySerSer 12       |
| DB |  | 579  | ATCATGCGTAAGCTGCACACTGCAATAATTGTGAGGCTGAGATACITTTTTTCTACTCCAGT 63    |
| QY |  | 130  | GlyGluLysIleAspGluValTyrLeuAsnLeuValLeuAspTyrValProGluThrVal 14      |
| DB |  | 639  | GCGCAGAAGAAAGACGAGCTTTTACCCTAAATCTGGTGTGGAAATATGTGCCCGCAGACAGTG 69   |
| QY |  | 150  | TyrArgValAlaArgHisTyrSerArgAlaLysAlinThrLeuProValIleTyrValLys 16     |
| DB |  | 699  | TACCGGTGGCCCGCCACTTCAACCAAGCCCAAGTTGACCTCCCTATCTCTATGTCCAG 75        |
| QY |  | 170  | LeuTyrMetTyrGlnLeuPheArgSerLeuAlaTyrIleHisSerPheGlyIleCysHis 18      |
| DB |  | 759  | GTGTACATGTACCACTCTTCCGCGAGCTTGGCTTACATCCACTCCAGGGCGTGTGTCCAC 81      |
| QY |  | 190  | ArgAspIleLysProGlnAsnLeuLeuLeuAspProAspThrAlaValLeuLysLeuCys 20      |
| DB |  | 819  | CSCGACATCAAGCCCAGAACCTGTCTGTGGAGCCCTGACACACTGTCTCTCAAGCTCTGC 87      |
| QY |  | 210  | AspPheGlySerAlaLysGlnLeuValArgIleGluProAsnValSerTyrIleCysSer 22      |
| DB |  | 879  | CAITTTTGGCAGTGCMAAGAGTGTGGTCCGAGGGGAGCCCAATGTCTCTACATCTGTGTTCT 93    |
| QY |  | 230  | ArgTyrTyrArgAlaProGluLeuIlePheCylAlaThrAspTyrThrSerSerileAsp 24      |
| DB |  | 939  | CGCTACTACCCGGGCCCCAGAGCTCATCTTTGGAGCCACCTGATTACACTCATCATCATCGAT 99   |
| QY |  | 250  | ValTrpSerAlaGlyCysValLeuAlaGluLeuLeuGlyGlnProIlePheProGly 26         |
| DB |  | 999  | GTTTGGTTCAGCTGGCTGTGTGTATGGCAGAGCTCTCTTGGGCGAGCCCATCTTCCTCTGGG 10    |
| QY |  | 270  | AspSerGlyValAspGlnLeuValGluIleLysValLeuGlyThrProThrArgGlu 28         |
| DB |  | 1059 | GACACTGGGTGGACCTGGTGGAGATCATCAAGTCTGGGAAACCCACACCCCGGAA 11           |

5

QY 290 GlnIleArgGluMetAspProAnThrThrGluPheProGlnIleLysAlaHis 309  
 Db 1119 CAAATCCGAGATGACCCCACTACAGGAGTTCAAGTTCCCTCAGATTAAAGCTCAC 1178

QY 310 ProTrpThrLysValPheArgProArgThrProGluAlaIleAlaLeuCySerArg 329  
 Db 1179 CCCTGGCAAAAGGTGTTCAATCTCGAACCGCGCAGAGCCATCGCGCTCTCTCTAGC 1238

QY 330 LeuLeuGluThrThrProThrAlaArgLeuThrProLeuGluAlaCysAlaHisSerPhe 349  
 Db 1239 CTGCTGGAGTACACCCCATCTCAAGGCTCTCCCACTAGAGGCGCTGTGGGCACAGCTTC 1298

QY 350 PheAspGluLeuArgAspProAsnValLysHisProAsnGlyArgAspThrProAlaLeu 369  
 Db 1299 TTTGATGACTGGGATGCTGGGAACCCAGCTGCTTAACACCGCCCACTTCCCTCTCTC 1358

QY 370 PheAsnPheThrThrGlnGluLeuSerSerAsnProProLeuAlaThrIleLeuLeuPro 389  
 Db 1359 TTCAACTTCAGCTGGTGGTGAACCTCTCATCCACCGCTCTCTCAACGCCATTTCTATCCCT 1418

QY 390 ProHisAlaArg 393  
 Db 1419 CCTCACTTGAGG 1430

RESULT 6  
 CR593340  
 LOCUS  
 DEFINITION full-length cDNA clone CS0DK005YF06 of Hela cells Cot 25-normalized  
 of Homo sapiens (human).  
 ACCESSION CR593340  
 VERSION CR593340.1 GI:50474147  
 KEYWORDS HTC; CNSUT\_CDNA.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 2124)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue  
 2 (bases 1 to 2124)  
 REFERENCE  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by life technologies, a  
 division of Invitrogen.  
 FEATURES  
 \* source  
 1. .2124  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DK005YF06"

/tissue\_type="Hela cells Cot 25-normalized"  
 /plasmid="pCMVSPORT\_6"

## ORIGIN

## Alignment Scores:

Pred. No.: 6,51e-166 Length: 2124  
 Score: 1609.00 Matches: 319  
 Percent Similarity: 77.3% Conservative: 24  
 Best Local Similarity: 71.8% Mismatches: 43  
 Query Match: 77.3% Indels: 58  
 DB: 6 Gaps: 4

US-10-733-816-2 (1-394) x CR593340 (1-2124)

QY 5 ProMetGluGlyGly---GlyMetSerGlyArgProArgThrThrSerPheAlaGluSer 23  
 Db 118 CCTTCGGGAGGCGGCCCTGGGGGCTCGGGCAGGGCGGAGCTAGCTCTTCGGGGAG--- 174

QY 24 CysLysPro-----ValGlnGln 20  
 Db 175 -----CCCGCGCGGAGGCGGAGGCGGGCGGCGCGGAGGCTCGGGCTCCGGC 228

QY 30 ProSerAlaPheGlySerMetLys----- 37  
 Db 229 CCAGGCGGCACCGCGCGCGGAAGGCATCTGTGGGGCCATGGGTGGGGGCGTGGGGGCC 288

QY 37 ----- 37  
 Db 289 TCGAGCTCCGGGGTGGACCCCGCGCGCAGCGCGGAGGAGCGAGCGCGCGCGCGCA 345

QY 38 -----ValSerArgAspLysAspGlySerLysValThrThr 49  
 Db 349 GGCACACTAGCTTCCCGCCCGCGGGTGAAGCTGGGCGGTGACAGCGGGAAGGTGACCA 408

QY 50 ValValAlaThrProGlyGlnGlyProAspArgProGlnGluValSerThrAspThr 63  
 Db 409 GTCTGAGCCACTCTAGGCCAAGGCCCGCAGAGCGCTCCCAAGAGTGGCTTACACGGACATC 468

QY 70 LysValIleGlyAsnGlySerPheGlyValValTyrglnAlaLysLeuCyAspSerGly 89  
 Db 469 AAAGTGATTGGCAATGGCTCATTTGGGGTCTGTACCAAGCAGCGCTGGCAGAGACAGG 528

QY 90 GluLeuValAlaIleLysLysValLeuGlnAspLysArgPheLysAsnArgGluLeuGln 109  
 Db 529 GAACCTAGTCGCCATCAAGAGAGTTTCCAGGACAAAGAGTTCAAGAACCGAGAGCTGCAG 588

QY 110 IleMetArgLysLeuAspHisCysAsnIleValArgLeuArgTyrgPhePheTyrsSer 129  
 Db 589 ATCATGGCTGAAGCTGGACCACTGCAATAATGTGAGGGCTGAGACTTTTCTACTCCAGT 640

QY 130 GlyGluLysLysAspGluValTyrglnLeuAsnLeuValLeuAspTyrgValProGluThrVal 149  
 Db 649 GCGGAGAGAAAGACGAGCTTTACCTTAATCTGGTGTGGATATGTGCCCGAGACAGTG 708

QY 150 TyrArgValAlaAlaArgHisTyrgSerArgAlaLysGlnThrLeuProValIleTyrgLys 169  
 Db 709 TACCGGGTGGCGGCCACTTCACCAAGGCCAAGTTGACCATCCCTATCTCTATGTCAG 768

QY 170 LeuTyrgMetTyrglnLeuPheArgSerLeuAlaTyrgIleHisSerPheGlyIleCysHis 189  
 Db 769 GTGTACATGTACCAGCTCTTCGGCAGCTTGGCGCTACATCATCCATCCCGAGGGCGTGTGT 828



|            |     |            |   |                 |
|------------|-----|------------|---|-----------------|
|            | -db | 363        | ATCCCTATCCTTACTGTCAAGGTTACATGTACACAGCTCTTCGCAGCGCTTGCGCTACATC     | 422             |
| Qy         |     | 183        | HISerPheGlyIleCyehisArgaspIleyProGlnAenLeuLeuAasProAsp            | 202             |
| Dd         |     | 423        | CACGCCAGGGCGTGTGTACCGCGCATCAAGCCCCAGAACCTCTGCTGGTGGACCTGCAC       | 482             |
| Qy         |     | 203        | ThrAlaValLeuLysLeuCysAspPheGlySerAlaIayGlnLeuValArgGlyGluPro      | 222             |
| Dd         |     | 483        | ACTGCTGTCTCAAGCTCTCGCGATTTCGGCAGTGCAAAGCAGTGGTCCAGGGGAGGCC        | 542             |
| Qy         |     | 223        | AsnValSerTyrlleCysaserArgTyTyTrArgAlaProGluLeuIlePheGlyAlaThr     | 242             |
| Dd         |     | 543        | ATGTCTCTCAATCTGTCTTCGTACTACCGCGCCCAGAGCTCACTTTTGGAGCCACT          | 602             |
| Qy         |     | 243        | AspTyThrSerSeriLeaspValTpSerAlaGlyCysValLeuAlaGluLeuLeu           | 262             |
| Dd         |     | 603        | GATTACACTCATCCATCGATGTTTGTGTGCTGCTGTGTACTTGGCAGAGCTCCTCTTG        | 662             |
| Qy         |     | 263        | GlyGlnProIlePheProGlyYaspSerGlyValaspGlnLeuValGluIleIleLysVal     | 282             |
| Dd         |     | 663        | GGCCAGCCCATCTTCCCTGGGGCAGTGGGGTGGACCAGCTGTGGAGATCATCAAGGTG        | 722             |
| Qy         |     | 283        | LeuGlyThrProThrArgGluGlnIleAtsGluMetAsnProAsnTyThrGluPheLys       | 302             |
| Dd         |     | 723        | CTGGGAACHCACACCCGGGACAATAATCCAGAGATGAACCCCACTACACGGAGTTCAAG       | 782             |
| Qy         |     | 303        | PheProGlnIleLysAlaHisProTyThrTyLysValPheArgProArgThrProProGlu     | 322             |
| Dd         |     | 783        | TTCCTTCAGATTAAGCTCACCCCTGGCANAGGTGTTCAATCTCGAAGCGCCGACAG          | 842             |
| Qy         |     | 323        | AlaIleAlaLeuCysSerArgLeuLeuGluTyThrProThrAlaArgLeuThrProLeu       | 342             |
| Dd         |     | 843        | GCCATCGGGCTCTGCTTAGCTCGAGTACACCCCATCTCTCAAGGCTCTCCCCACTA          | 902             |
| Qy         |     | 343        | GluAlaCysAlaHisSerPhePheAspGluLeuArgAspProAsnValLysHisProAsn      | 362             |
| Dd         |     | 903        | GAGGCTCTGGCGACAGCTTCTTGTATGAACATGGCATGTCTGGGAACCCAGCTGCCTAAC      | 962             |
| Qy         |     | 363        | GlyArgAspThrProAlaLeuPheAsnPhrThrGlnGluLeuSerSerAsnProPro         | 382             |
| Dd         |     | 963        | AACCGCCACTTCCCTCTTTCNACTGCTGGTGACTCTCCATCCCAACCGCTCT              | 1022            |
| Qy         |     | 383        | LeuIleLeuLeuProProHisAlaArg                                       | 393             |
| Dd         |     | 1023       | CTCAAGCCACTTCTATCCCTCCTCACTTGAGG                                  | 1055            |
| RESULT 8   |     |            |   |                 |
| CR749620   |     |            |   |                 |
| LOCUS      |     | CR749620   | 3536 bp mRNA linear   | HTC 19-AUG-2004 |
| DEFINITION |     |            | Hom sapiens mRNA; cDNA DKFP668D0638 (from clone DKFP668D0638).    |                 |
| ACCESSION  |     | CR749620   |   |                 |
| VERSION    |     | CR749620.1 | GI:51476895   |                 |
| KEYWORDS   |     |            | HTC.  |                 |
| SOURCE     |     |            | Homo sapiens (human)  |                 |
| ORGANISM   |     |            | Homo sapiens  |                 |
| REFERENCE  |     |            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |                 |
| AUTHORS    |     |            | Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;       |                 |
|            |     |            | Hominidae; Homo.  |                 |
|            |     |            | 1 (bases 1 to 3536)   |                 |
|            |     |            | Ottenwaelder,B., Obermaier,B., Deutschenbaub,S., Schaipp,A.,      |                 |
|            |     |            | Mewes,H.W., Weill,B., Anid,C., Osanger A., Fobo,G., Han,M. and    |                 |



|            |   |  |                 |
|------------|---|--|-----------------|
| QY         | 311   | TroThrLysValPheArgProArgThrProGluAlaIleAlaLeuCysSerArgLeu    | 330             |
| Db         | 866   | TGGACAAAGGTTTCAAATCTCGACAGCGCCGACAGGCCATCGGGCTGTGCTCTACCGTG  | 925             |
| QY         | 331   | LeuGluTyrThrProThrAlaArgLeuThrProLeuGluIaCysAlaHisSerPhePhe  | 350             |
| Db         | 936   | CTGGAGTACACCCATCCTCAAGGCTCTCCCACTAGAGGCTGTGCGCACAGCTCTTT     | 985             |
| QY         | 351   | AspGluLeuArgAspProAsnValLysHisProAsnGlyArgAspThrProAlaLeuPhe | 370             |
| Db         | 986   | GATGAATGCGATGTCTGGGAACCCAGCTGGCTTAAACAACGCCCACTTCCCCCTCTCTTC | 1045            |
| QY         | 371   | AsnPheThrThrGlnGluLeuSerSerAsnProProLeu                      | 383             |
| Db         | 1046  | AACCTCAGTGTGGTGAAGTCTCCATCCAACCGTCTCTA                       | 1084            |
| RESULT 10  |   |  |                 |
| LOCUS      | CX795595  | 841 bp mRNA linear   | EST 31-JAN-2005 |
| DEFINITION | JGI_CAAJ11693.fwd NIH_XGC_tropBrr2 Xenopus tropicalis cDNA clone  |  |                 |
| ACCESSION  | IMAGE:7641717 5', mRNA sequence.  |  |                 |
| VERSION    | CX795595  |  |                 |
| KEYWORDS   | CX795595.1 GI:58350224  |  |                 |
| SOURCE     | EST.  |  |                 |
| ORGANISM   | Xenopus tropicalis (western clawed frog)  |  |                 |
| REFERENCE  | Xenopus tropicalis  |  |                 |
| AUTHORS    | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.   |  |                 |
| TITLE      | 1 (bases 1 to 841)  |  |                 |
| JOURNAL    | Richardson,P., Lucas,S., Rokhsar,D., Dettter,J.C., Ng,D.C., Brokstein,P. and Lindquist,E.A.   |  |                 |
| COMMENT    | DOE Joint Genome Institute Xenopus tropicalis EST project Unpublished (2004)  |  |                 |
|            | Other_ESTs: JGI_CAAJ11693.rev   |  |                 |
|            | Contact: Lindquist,E.A., Richardson,P.  |  |                 |
|            | DOE Joint Genome Institute  |  |                 |
|            | 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  |  |                 |
|            | Tel: 925 296 5600   |  |                 |
|            | Fax: 925 296 5710   |  |                 |
|            | Email: cdna@jgi-psf.org   |  |                 |
|            | Tissue Procurement: Timothy Grammer (Richard M. Harland Laboratory, University of California, Berkeley)   |  |                 |
|            | http://tropicalis.berkeley.edu/home/  |  |                 |
|            | cDNA Library Preparation: DOE Joint Genome Institute:   |  |                 |
|            | http://www.jgi.doe.gov  |  |                 |
|            | DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov  |  |                 |
|            | Clone Distribution: I.M.A.G.E. Consortium/ILLNL:  |  |                 |
|            | http://image.llnl.gov   |  |                 |
|            | Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert. |  |                 |
|            | Plate: CNAJ 0121 row: j column: 19  |  |                 |
|            | High quality sequence stop: 794.  |  |                 |
| FEATURES   | Location/Qualifiers   |  |                 |
| SOURCE     | 1..841  |  |                 |
|            | /organism="Xenopus tropicalis"  |  |                 |
|            | /mol_type="mRNA"  |  |                 |
|            | /db_xref="taxon:8364"   |  |                 |
|            | /clone="IMAGE:7641717"  |  |                 |

```

/tissue_type="Brain"
/dev_stage="Adult"
/lab_host="ElectroMAX DH10B"
/clone_lib="NIH_XGC_tropBn2"
/notes="Vector: pCMVSPORT6; Site: 1: SalI; Site 2: NotI;
This library was made from dt primed cDNA and cloned into
Invitrogen pCMVSPORT6 vector. The work was done at DOE
Joint Genome Institute. Poly A RNA was primed with 5'
GACTAGTCTAGATCGGAG CGGCCGCTTTT TTTT 3'. cDNA
were ligated to SalI adapter (5' TCGACCCACGCGCCG and
5'CGACCGCGGG), digested with NotI, size fractionated in
1.1% agarose gel electrophoresis and ligated into NotI and
SalI digested pCMVSPORT6 vector."

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 1.68e-148 Length: 841
Score: 1446.00 Matches: 274
Percent Similarity: 98.9% Conservative: 3
Best Local Similarity: 97.9% Mismatches: 3
Query Match: 69.5% Indels: 0
DB: 9 Gaps: 0

```

US-10-733-816-2 (1-394) x CX795595 (1-841)

```

QY 42 LysAspGlySerIysValThrValValAlaThrProGlyGlnGlyProAspArgPro 61
DB 2 AAAGATGCCAGCAAGTACACAGTAGTAGCGACCCCGCGGCGCGCGCGCGAG 61
QY 62 GlnGluValSerIysThrAspThrIysValIleGlyAsnGlySerPheGlyValValTyr 81
DB 62 CAAGAAGTCAGTCACCGCACCAAGGTTATTGGCAAGCGGTCTTTGGGGTTGGTAC 121
QY 82 GlnAlaIysLeuCysAspSerGlyGluLeuValAlaIleIysValLeuGlnAspIys 101
DB 122 CAGGCCAAACTGTGCGACACCGGGAGCTGGTGCTATCAGAAGGTGCTACAGCAAG 181
QY 102 ArgPheIysAsnArgGluLeuGlnIleMetArgIysLeuAspHisCysAsnIleValArg 121
DB 182 AGATTTAAGAACCGAGAACTGCAGATCATGAGAAAACCTGGATCACTGCACACTTGTGCGT 241
QY 122 LeuArgTyrPhePheTyrSerSerGlyGluIysIysAspGluValTyrLeuAsnLeuVal 141
DB 242 CTGGCGTACTTCTTCTACTCCAGCGGGAGAGAGAGCAAGGTGTACTCAACTTGGTG 301
QY 142 LeuAspTyrValProGluThrValTyrArgValAlaArgHisTyrSerArgAlaIysGln 161
DB 302 CTGGACTACGTTCCCGAACTGTGTACAGAGTGGCCAGACATTACAGCCGGGCTAAGCAG 361
QY 162 ThrLeuProValIleTyrValIysLeuTyrMetTyrGlnLeuPheArgSerLeuAlaTyr 181
DB 362 GCGTCCCGATGATCTATGTCAGTTGTACATGTACAGCTGTTCGGCAGTTTAGCCTAC 421
QY 182 IleHisSerPheGlyIleCysHisArgAspIleIysProGlnAsnLeuLeuLeuAspPro 201
DB 422 ATCCATTCTTTGGGATCTGCCCGGGATATCAACACAGAACCTTCTGCTGGACCA 481
QY 202 AspThrAlaValLeuIysLeuCysAspPheGlySerAlaIysGlnLeuValArgGlyGlu 221
DB 482 GAAACTGCTGCTACTCAAGCTGTGTGCTTTGGCAGTGCACAGCAGCTGGTACCGGAG 541

```

```

QY 222 ProAsnValSerTyrIleCysSerArgTyrArgAlaProGluLeuIlePheGlyAla 241
DB 542 CCCAACGTGCTCTACATCTGCTCCGCTACTACAGGCCGCCAGAGTTGATCTTTGGAGCC 601
QY 242 ThrAspTyrThrSerSerIleAspValTrpSerAlaGlyCysValLeuAlaGluLeuLeu 261
DB 602 ACCGATTACACTTCCAGCTATTACGCTGCTGTCAGCAGGCTGCGTATTGGCGGAGCTGCTG 661
QY 262 LeuGlnPheProIlePheProGlyAspSerGlyValAspGluLeuValGluIleIleIys 281
DB 662 CTGGGTGAGCCCATCTTCTTGGGGACAGCGCGTGGATCAGCTGGTTGAGATAATCAAG 721
QY 282 ValLeuGlyThrProThrArgGluGlnIleArgGluMetAsnProAsnTyrThrGluPhe 301
DB 722 GTTCTGGGTACCCCGCTCTCGAGAGCAGATCCGAGAAATGAAACCCCACTACTGTAATTC 781
QY 302 LysPheProGlnIleIysAlaHisProTrpThrIysValPheArgProAsnTyrProPro 321
DB 782 AAGTTCCACAGATAAGGCGCATCTTGGCAAAAGGTGTTCCGGGGCGCAGCCCTCC 841
RESULT 11
DN530922/c 818 bp mRNA linear EST 11-MAR-2005
LOCUS 1358786 MARC 7BOV Bos taurus cDNA 3', mRNA sequence.
DEFINITION DN530922
ACCESSION DN530922
VERSION DN530922.1 GI:60976699
KEYWORDS EST.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 818)
AUTHORS Smith,T.P.L., Roberts,A.J., Echterkamp,S.E., Chitko-McKown,C.G.,
Wray,J.E. and Keefe,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: RLK8053 row: M column: 14
Seq primer: TAGAGGCACAGTCGAGG.

```

## FEATURES

```

source
1..818
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 7BOV"
/notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including ovary, hindbrain, uterus, and day-30 whole
embryos."

```

## ORIGIN

Alignment Scores:  
 Pred. No.: 4.28e-145 Length: 818  
 Score: 1415.00 Matches: 271  
 Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 68.0% Indels: 0  
 DB: 9 Gaps: 0

US-10-733-816-2 (1-394) x DN530922 (1-818)

QY 35 SerMetLysValSerArgAspLysAspGlySerLysValThrThrValValAlaThrPro 54  
 |||||  
 DB 816 AGCATGAAGTTAGCAGAGACAAAGATGCGCAGCAGGTGACCAAGTGTGGCACTCCT 757  
 |||||

QY 55 GlyGlnGlyProAspArgProGlnGluValSerTyrThrAspThrLysValIleGlyAsn 74  
 |||||  
 DB 756 GGGCAGGGTCCAGACAGCCCAAGCAAGTCAAGTATACAGACACTAAAGTGTGCAAT 697  
 |||||

QY 75 GlySerPheGlyValValTyrGlnAlaLysLeuCysAspSerGlyGluLeuValAlaIle 94  
 |||||  
 DB 696 GGGTCATTTGGTTCGTGTATCAAGCCAAACTTTGTGATTCAGGAGAACTGGTGGCAATC 637  
 |||||

QY 95 LysLysValLeuGlnAspLysArgPheLysAsnArgGluLeuGlnIleMetArgLysLeu 114  
 |||||  
 DB 636 AAGAAGTATTGCAAGACAGACAGATTTAAGAACCGAGAGCTCAAGTCAAGTGAAGAAAGAT 577  
 |||||

QY 115 AspHisCysAsnIleValArgLeuArgTyrPheTyrSerSerGlyGluLysLysAsp 134  
 |||||  
 DB 576 GATCACTGTAACATAGTCCGATTGCGTTATTTCTTACTCAAGTGGTGAAGAAAGAT 517  
 |||||

QY 135 GluValTyrLeuAsnLeuValLeuAspTyrValProGluThrValTyrArgValAlaArg 154  
 |||||  
 DB 516 GAGGTCTATCTTAATCTGCTGGACTACGTTCCGAAACAGCTACAGAGTCGCCAGA 457  
 |||||

QY 155 HisTyrSerArgAlaLysGlnThrLeuProValIleTyrValLysLeuTyrMetTyrGln 174  
 |||||  
 DB 456 CACTATAGTCGAGCCAAACAGACGCTCCCTGTGATCTATGTCAAGTTGTATATGTATCAG 397  
 |||||

QY 175 LeuPheArgSerLeuAlaTyrIleHisSerPheGlyIleCysHisArgAspIleLysPro 194  
 |||||  
 DB 396 CTGTTCCGAAGTTTACCTATATCCATTCCTTTGGAACTCTGCCATCGCGATATTAAACA 337  
 |||||

QY 195 GlnAsnLeuLeuAspProAspThrAlaValLeuLysLeuCysAspPheGlySerAla 214  
 |||||  
 DB 336 CAGAACTCTTGTGGATCCTGATACAGCTGCTTGAACCTCTGTGACTTTGGAAGTGA 277  
 |||||

QY 215 LysGlnLeuValArgGlyGluProAsnValSerTyrIleCysSerArgTyrTyrArgAla 234  
 |||||  
 DB 276 AAGCAGCTGTCGCTGGAGAACCCCAATGTTTCGTATATCTCTGTTCTCGTACTATAGGCA 217  
 |||||

QY 235 ProGluLeuIlePheGlyAlaThrAspTyrThrSerSerIleAspValTyrSerAlaGly 254  
 |||||  
 DB 216 CCAGAGTTGATCTTTGGAGCCACTGATATATACCTCTAGTATAGATGTATGGTCTGCAGGC 157  
 |||||

QY 255 CysValLeuAlaGluLeuLeuGlyGlnProIlePheProGlyAspSerGlyValAsp 274  
 |||||  
 DB 156 TGTGTGTGGCTGAGCTGTCTAGCAGACCCATATTTCCAGGAGACAGCTGTGTGGAT 97  
 |||||

QY 275 GlnLeuValGluIleLysValLeuGlyThrProThrArgGluGlnIleArgGluMet 294  
 |||||

96 CAGTTGGTGAATAATCAAGTCTCTGGGACACCAACAGGGAGCAAAATTCGAGAAATG 37

295 AenProAsnTyrThrGluPheLysPheProGln 305  
 |||||

36 AACCCAAACTACACGGAATTCAAATTCCTCA 4

RESULT 12  
 CO579074

LOCUS CO579074 993 bp mRNA linear EST 20-JUL-2004  
 DEFINITION ILLUMIGEN\_MQ\_50668 Katze\_MMIL\_Macaca mulatta cDNA clone  
 IBIUM.16936 5' similar to Bases 5 to 992 highly similar to human  
 GSK3A (Hs.435970), mRNA sequence.

ACCESSION CO579074  
 VERSION CO579074.1 GI:50409944  
 KEYWORDS EST.

SOURCE Macaca mulatta (rhesus monkey)  
 ORGANISM Macaca mulatta  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Cercopithecoidea; Cercopithecinae; Macaca.

REFERENCE 1 (bases 1 to 993)  
 MAGNESS,C.L., Feilun,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,  
 Prohl,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and  
 Iadonato,S.P.  
 Analysis of the Macaca mulatta transcriptome and the sequence  
 divergence between Macaca and human  
 Genome Biol. 6 (7), R60 (2005)  
 15998449

JOURNAL PUBMED  
 COMMENT Contact: C. Magness  
 Illumigen Biosciences Inc.  
 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA  
 Tel: 2063780400  
 Fax: 2063780408  
 Email: cmagness@illumigen.com  
 Sequenced on 2004.07.16. 702 020 bases. Library Preparation: Prcf.  
 Michael Katze Lab at University of Washington DNA Sequencing.  
 Illumigen Biosciences Inc. For further information, see  
 http://www.macaque.org

PCR Primers  
 FORWARD: CCTCACTAAAGGCAACAAA  
 BACKWARD: CACTATAGGCGAATTGGTA  
 Insert Length: 993 Std Error: 0.00  
 Plate: CL000422 row: E column: 09  
 Seq primer: CCTCACTAAAGGCAACAAA  
 POLYA=No.

FEATURES  
 Location/Qualifiers  
 1..993  
 /organism="Macaca mulatta"  
 /mol\_type="mRNA"  
 /strain="Indian"  
 /db\_xref="taxon:9544"  
 /clone="IBIUM.16936"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="Electromax DH10B"  
 /clone\_lib="Katze\_MMIL"  
 /notes="Organ: ileum; Vector: pDONR 222; Site 1: BarG I;  
 Site 2: BarG I; Created from Clonewiner cDNA Library  
 Construction kit (catalog #18249-029)"

ORIGIN



## Alignment Scores:

Pred. No.: 1-78e-144 Length: 993  
 Score: 1410.50 Matches: 274  
 Percent Similarity: 88.6% Conservative: 22  
 Best Local Similarity: 82.0% Mismatches: 25  
 Query Match: 67.8% Indels: 13  
 DB: 8 Gaps: 2

US-10-733-816-2 (1-394) x C0579074 (1-993)

QY 52 AlaThrProGlnGlnGlyProAspArgProGlnGlnValSerTyrThrAspThrLysVal 71  
 DB 5 GCCACTCTAGGCCAAGGTCACAGAGCGCTCCCAAGAGGTGCTTACACAGACATCAAAAGTG 64  
 QY 72 IleGlyAsnGlySerPheGlyValValTyrGlnAlaLysLeuCysAspSerGlyGluLeu 91  
 DB 65 ATGGCAATGGCTATTGGGGTGGTGTACCAAGCAGCGCTGGCAGACACCGGAACGTG 124  
 QY 92 ValAlaIleLysValLeuGlnAspLysArgPheLysAsnArgGluLeuGlnIleMet 111  
 DB 125 GTCCGCATCAAGAAAGGTTCTCCAGGACAAGAGGTTCAAGAACCGAGAGCTGCAGATCATG 184  
 QY 112 ArgLysLeuAspHisCysAsnIleValArgLeuArgTyrPhePheTyrSerSerGlyGlu 131  
 DB 185 CGTAAGCTGGACCACTGCATATATTGTGAGCTGAGACTTTTCTACTCCAGTGGGAG 244  
 QY 132 LysLysAspGluValTyrLeuAsnLeuValLeuAspTyrValProGluThrValTyrArg 151  
 DB 245 AAGAAGATGAGCTTTACCTAATCTGGTGTGGAATATGTGCCGAGACAGTGTACCGG 304  
 QY 152 ValAlaArgHisTyrSerArgAlaLysGlnThrLeuProValIleTyrValLysLeuTyr 171  
 DB 305 GTGGCCGCCACTTACCAGAGCCAAAGTTGACCATCTCTATCTATGCAAGGTGATC 364  
 QY 172 MetTyrGlnLeuPheArgSerLeuAlaTyrIleHisSerPheGlyLeCysHisArgAsp 191  
 DB 365 ATGTACCAGCTCTTCGAGCTTGGCTACATCCACTCCAGGCGGTGTCTACCGTGAC 424  
 QY 192 IleLysProGlnAsnLeuLeuAspProAspThrAlaValLeuLysLeuCysAspPhe 211  
 DB 425 ATCAAGCCCCAGAACCTCTGCTGGACCTTGACACTCTCTCTCACTCTCGGATTT 484  
 QY 212 GlySerAlaLysGlnLeuValArgGlyGluProAsnValSerTyrIleCysSerArgTyr 231  
 DB 485 GCACGTGGAAGCAGTGTGTCGAGGAGGCCAAATGTCTCTACATCTGTTCTCGCTAC 544  
 QY 232 TyrArgAlaProGluLeuIlePheGlyAlaThrAspTyrThrSerIleAspValTyr 251  
 DB 545 TACCGGCCCCAGAGCTCATCTTTGGAGCCACTGATTACACCTCATCTCCATCGATGTTGG 604  
 QY 252 SerAlaGlyCysValLeuAlaGluLeuLeuGlyGlnProIlePheProGlyAspSer 271  
 DB 605 TCAGTGGCTGCTGCTGCGGAGCTCTCTCCGCGCAGGCCCATTTCTCTGGGAGACGT 664  
 QY 272 GlyValAspGlnLeuValIleLysValLeuGlyThrProThrArgGluGlnIle 291  
 DB 665 GGGGTGACCAAGCTGCTGGAGATCATCAAGGTGCTGGACACCAACCCGGGCAACATC 724  
 QY 292 ArgGluMetAsnProAsnTyrThrGluPheLysPheProGlnIleLysAlaHisProTyr 311  
 DB 725 CGAGAGATGAACCCAACTACACGAGTTCATGTTCTCCATTAAGATCAACCTCCCTGG 784

QY 312 ThrLysValPheArgProArgThrProGluAlaIleAlaLeuCysSerArgLeuLeu 321  
 DB 785 ACRAAGGTTCATATCTCGAAGCCGCCAGAGCATCCGCTCTGCTCTAACCCTGCTG 844  
 QY 332 GlutYrThrProThrAlaArgLeu-ThrProLeuGluAlaCysAlaHisSerPhe-PheA 351  
 DB 845 GAGTACCCCGCCGCTCAAGGCTCCGCCCTTTAAAGAGCTGCGCACACAGCTTCCTTTG 904  
 QY 351 sPGLuLeuArgAspProAsnValLysHisProAsnGlyArgAspThrProAlaLeuPheA 371  
 DB 905 AAGAAGCTGCAAGCGCGGG-----AACCCAGCTCCCTTAAAAAACCCCCCACTTC- 937  
 QY 371 snPheThrThrGlnGluLeuSerSerAsnProProLeu 383  
 DB 958 -----CCCCCCCCCTT 969

## RESULT 13

DT309300

LOCUS DT309300

DEFINITION

males and females pooled (L) Pimephales promelas cDNA clone

CAAX1138 5', mRNA sequence.

ACCESSION DT309300

VERSION DT309300.1 GI:73649963

KEYWORDS EST.

SOURCE Pimephales promelas

ORGANISM Pimephales promelas

REFERENCE 1 (bases 1 to 869)

AUTHORS Richardson, P., Lucas, S., Rokhsar, D., Dettler, J.C., Ng, D.C.,

Brokstein, P. and Lindquist, E.A.

TITLE DOE Joint Genome Institute Pimephales promelas EST project

JOURNAL Unpublished (2005)

COMMENT Other\_ESTs: JGI\_CAAX1138.rev

Contact: Lindquist, E.A., Richardson, P.

DOE Joint Genome Institute

2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Tel: 925 296 5600

Fax: 925 296 5710

Email: cdna@jgi-psf.org

cDNA Library Preparation: DOE Joint Genome Institute:

http://www.jgi.doe.gov

DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov

Naming Conventions: EST name is generated by the concatenation of

the JGI Clone id and the direction of sequencing. The suffix '.fwd'

indicates a forward sequencing read of the insert. It does not

necessarily reflect the orientation of the insert.

Plate: CAAX 0009 row: d column: 22

High quality sequence stop: 788.

Location/Qualifiers

1. .869

/organism="Pimephales promelas"

/mol\_type="mRNA"

/db\_xref="taxon:90988"

/clone="CAAX1138"

/tissue\_type="testis"

/clone\_lib="CAAX Pimephales promelas testis 7-8 month

adults, males and females pooled (L)"

## FEATURES

source

/note="Vector: pCMVSPORT6; The library was made from dT primed cDNA and cloned into invitrogen vector pCMVSPORT6. Poly A RNA were primed with an oligo dT primer (5'-GACTAGTCTTAGATCGGAGCGCGCCCTTTTITTTT -3') ligated to a SalI adapter (5'-TCGACCCACCGCTCGG and 5'-CGACCGCTGGG) and digested with NotI. cDNA was size selected using 1.1% agarose gel electrophoresis (L -0.5-1.2k, M -1.2-2.5k, H ->2.5k) then ligated into NotI and SalI digested pCMVSPORT6 vector. The work was done at DOE Joint Genome Institute."

## ORIGIN

Alignment Scores:  
 Pred. No.: 9,88e-144 Length: 869  
 Score: 1403.00 Matches: 258  
 Percent Similarity: 96.5% Conservative: 20  
 Best Local Similarity: 89.6% Mismatches: 10  
 Query Match: 67.4% Indels: 0  
 DB: 10 Gaps: 0

US-10-733-816-2 (1-394) x DT039300 (1-869)

QY 80 ValTyrGlnAlaLysLeuCyAspSerGlyGluLeuValAlaLysValLeuGln 99  
 DB 3 GTGTACAGGCTGCTATCGACGAGGATGGTCCATTAAGAGGTGCTTCAG 62  
 QY 100 AspLysArgPheLysAsnArgGluLeuGlnLeuMetArgLysLeuAspHisCysAsnIle 119  
 DB 63 GATAACGGTTTAAAGACCGTGAGCTACAGATATGAGGAGTTCGACCACTGTATAT 122  
 QY 120 ValArgLeuArgTyrPhePheTyrSerSerGlyGluLysAspGluValTyrLeuAsn 139  
 DB 123 GTACGGCTAGCTACTTCTTCTACTCCAGTGAGAGAAAGATGAAGTGTATCTAAAT 182  
 QY 140 LeuValLeuAspTyrValProGluThrValTyrArgValAlaArgHisTyrSerArgala 159  
 DB 183 CTGGTGCTGGATTTGTTCCAGAACTGTGTACAGGCTGGCAGCCATTTCAACAAGTCC 242  
 QY 160 LysGlnThrLeuProValIleTyrValLysLeuTyrMetTyrGlnLeuPheArgSerLeu 179  
 DB 243 AAGACCACTCCCTATCTATGTCAGAGTATATGATATGATATTCGAGCTGTG 302  
 QY 180 AlaTyrIleHisSerPheGlyIleCysHisArgAspIleLysProGlnAsnLeuLeuLeu 199  
 DB 303 GCGTATATTCATTCGAGGCGCTGCCATGACATCAAGCCAGCAATCTCCTGGTG 362  
 QY 200 AspProAspThrAlaValLeuLysLeuCysAspPheGlySerAlaLysGlnLeuValArg 219  
 DB 363 GACCCAGACACAGCTGTACTCAAACTGTGTGACTTGGCAGTGCAAAACAGTTAGTTCT 422  
 QY 220 GlyGluProAsnValSerTyrIleCysSerArgTyrTyrArgAlaProGluLeuIlePhe 239  
 DB 423 GGGGAGGCCAATGTGTCAATACATCTGTTCGGGTATTTACCGGCTCTGAGCTTAATTC 482  
 QY 240 GlyAlaThrAspTyrThrSerIleAspValTyrSerAlaGlyCysValLeuAlaGlu 259  
 DB 483 GGAGCCAGGATTTACATCTCCATATGATATCTGTGACCGCTGTGTATAGCAGAG 542  
 QY 260 LeuLeuLeuGlyGlnProIlePheProGlyAspSerGlyValAspGlnLeuValGluIle 279  
 DB 543 CTGCTTTTGGGACAGCCATATTTCCCGGTGACAGTGGAGTGACCACTAGTGGAGATC 602

QY 280 IleLysValLeuGlyThrProThrArgGluGluGlnIleArgGluMetAsnProAsnTyrThr 293  
 DB 603 ATCAAGGTTTTGGGACCCCCACAGAGAACAGATCCGAGAGATGAACCCCAACTACACA 662  
 QY 300 GluPheLysPheProGlnIleLysAlaHisProTyrThrLysValPheArgProArgThr 319  
 DB 663 GAGTTTAAATTTCCACAGATCAAGCACACCTTTGGACAAAGGTGTTTACCGCAGGACC 722  
 QY 320 ProProGluAlaIleAlaLeuCySerArgLeuLeuGluTyrThrProThrAlaArgLeu 339  
 DB 723 CCTCTGAAGCCATCTCATTTCTCTCTGTGTGTACACGCGGTGACGCGGCTC 782  
 QY 340 ThrProLeuGluAlaCysAlaHisSerPhePheAspGluLeuArgAspProAsnValLys 359  
 DB 783 TCGCGGCTGGAGGCTGTGCACATGCTTCTTTGATGAGCTGCGCAGCCGACGACGT 842  
 QY 360 HisProAsnGlyArgAspThrPro 367  
 DB 843 CTGCTTAATGCGGAGAACTTCCC 866

## RESULT 14

AK164048 2190 bp mRNA linear HTC 21-SEP-2005  
 Mus musculus 7 days embryo whole body cDNA, RIKEN full-length  
 enriched library, clone:C430034M07 product:glycogen synthase kinase  
 3 beta, full insert sequence.

ACCESSION AK164048.1 GI:74208723

VERSION AK164048.1

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

## REFERENCE

1 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

10349636

## REFERENCE

2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)

11042159

## REFERENCE

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
 Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M.,  
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
 Yamamoto,K., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
 Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,  
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,  
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

11076861

## REFERENCE

4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,  
 11076861

- Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Cavaant, T., Fleishmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, T., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Waehio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, R., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
- RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium
- Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 11217851
- 5
- Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Ootani, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Bruscia, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
- PANTOM Consortium
- Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420 (6915), 563-573 (2002)  
 12466851
- 6
- Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E.,

- Ambesi-Impombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., Di Bernarado, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgi-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jaki, M., Kanapin, A., Katoh, M., Kawasawa, Y., Keisjo, J., Kitamura, H., Kitano, H., Kollas, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakachi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.P., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugiyama, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, Z., Zimmer, A., Hide, W., Bult, C., Grimond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori, K., Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Wachihi, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.
- PANTOM Consortium
- The transcriptional landscape of the mammalian genome  
 Science 309 (5740), 1559-1563 (2005)  
 16141072
- 7
- Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Fang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.
- RIKEN Genome Exploration Research Group
- Antisense transcription in the mammalian transcriptome  
 Science 309 (5740), 1564-1566 (2005)  
 16141073
- 8 (bases 1 to 2190)
- Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Wachihi, A., Muramatsu, M. and Hayashizaki, Y.
- Direct Submission
- Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa, 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,  
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

## COMMENT

CDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in Riken.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/  
Location/Qualifiers

## FEATURES

source  
1. .2190  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:C430034M07"  
/db\_xref="taxon:10090"  
/clone="C430034M07"  
/tissue\_type="whole body"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="7 days embryo"  
misc\_feature  
1. .2190  
/note="glycogen synthase kinase 3 beta (MGD|MG1:1861437  
GB|BC006936, evidence: BLASTN, 100%, match=788)  
putative"

## ORIGIN

Alignment Scores:  
Pred. No.: 2-22e-141 Length: 2190  
Score: 1387.00 Matches: 263  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 66.7% Indels: 0  
DB: 6 Gaps: 0  
US-10-733-816-2 (1-394) x AK164048 (1-2190)  
QY 52 AlaThrProGlyGlnGlyProAspArgProGlnGluValSerTyrThrAspThrLysVal 71  
|||||  
DB 1 GCAACTCTGCCCGAGGCTCTGACAGCCACAGGAGTCAATTATACAGACAGCAAGTG 60  
QY 72 IleGlyAsnGlySerPheGlyValValTyrGlnAlaLysLeuCysAspSerGlyGluLeu 91  
|||||  
DB 61 ATTGAAATGATCATTTGGTGTGTATATCAAGCCAACTTTGTGATTTCTGGAGAACTG 120  
QY 92 ValAlaLysLysValLeuGlnAspLysArgPheLysAsnArgGluLeuGlnIleMet 111  
|||||  
DB 121 GTTGCCATCAAGAAAGTTCTACAGGACAGCGATTTAAGAACGAGAGCTCCAGATCATG 180  
QY 112 ArgLysLeuAspHisCysAsnIleValArgLeuArgTyrPhePheTyrSerSerGlyGlu 131  
|||||  
DB 181 AGAAGCTAGACCACTGTAACTATAGTCCGACTGCGGTATTTCTTACTCGAGTGGTAG 240  
QY 132 LysLysAspGluValTyrLeuAsnLeuValLeuAspTyrValProGluThrValTyrArg 151  
|||||  
DB 241 AAGAAGATGAGGTCTACCTTAACCTGGTGGACTATGTCTCGAGACAGAGTGTACAGA 300  
QY 152 ValAlaArgHisTyrSerArgAlaLysGlnThrLeuProValIleTyrValLysLeuTyr 171  
|||||  
DB 301 GTCCGACAGACTATAGTCGAGCCAGCAGACACTCCCTGTGATCTATGTCAGTTGTAT 360

QY 172 MetTyrGlnLeuPheArgSerLeuAlaTyrIleHisSerPheGlyIleCysHisArgAsp 191  
|||||  
DB 361 ATGATATCAGCTGTTCAGAGCTAGCTATATCCATTCCTTTGGATCTGCCATCGAGAC 420  
QY 192 IleLysProGlnAsnLeuLeuAspProAspThrAlaValLeuLysLeuCysAspPhe 212  
|||||  
DB 421 ATTAACACACAGAACCTCTCTGTGGATCTCTATACAGCTGTATTAAAACTCTGTGACTT 480  
QY 212 GlySerAlaLysGlnLeuValArgGlyGluProAsnValSerTyrIleCysSerArgTyr 231  
|||||  
DB 481 GGAAGTCAAGCAGCTGCTCCGAGGAGAGCCCAATGTTTCATATATCTGTCTCCGTAC 540  
QY 232 TyrArgAlaProGluLeuIlePheGlyAlaThrAspTyrThrSerSerIleAspValTyr 251  
|||||  
DB 541 TACGGCACCAGAGTTGATCTTTGGAGCCACTGATTCACGTCCTCAGTATAGATGTATGG 600  
QY 252 SerAlaGlyCysValLeuAlaGluLeuLeuGlyGlnProIlePheProGlyAspSer 271  
|||||  
DB 601 TCTGCAGGCTGTGTGTGGCTGAATTGTTGCTAGGACACCAATATTTCTGGGGACAGT 660  
QY 272 GlyValAspGlnLeuValGluIleLysValLeuGlyThrProThrArgGluGlnIle 291  
|||||  
DB 661 CGTGTGGATCAGTTGTGTGGAATAATAAAGGTCTTAGGAAACCAACAGGAGCAAAAT 720  
QY 292 ArgGluMetAsnProAsnTyrThrGluPheLysPheProGlnIleLysAlaHisProTyr 311  
|||||  
DB 721 AGAGAAATGAACCCAAATTTATACAGAAATTCCTCAATCAAGGCACATCTCTTGG 780  
QY 312 ThrLysVal 314  
|||||  
DB 781 ACAAGGTG 789  
RESULT 15  
BO651422  
LOCUS BO651422  
DEFINITION BO651422 890 bp mRNA linear EST 15-JUL-2002  
AGENCOURT\_8341134 NIH\_MGC\_100 Homo sapiens cDNA clone IMAGE:6268349  
5', mRNA sequence.  
ACCESSION BO651422  
VERSION BO651422.1 GI:21775594  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 890)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapba-r@mail.nih.gov  
Tissue Procurement: CGAP (Stanford)  
cDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2440 row: c column: 06  
High quality sequence stop: 717.

# FEATURES

Location/Qualifiers

1. .890  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6268349"  
/tissue\_type="hepatocellular carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_100"  
/note="Organ: liver; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

### Alignment Scores:

Prod. No.: 4.57e-141 Length: 890  
Score: 1379.00 Matches: 262  
Percent Similarity: 94.9% Conservative: 17  
Best Local Similarity: 89.1% Mismatches: 13  
Query Match: 66.3% Indels: 2  
DB: 3 Gaps: 0

US-10-733-816-2 (1-394) x BQ651422 (1-890)

QY 60 ArgProGlnGluValSerTyrThrAspThrLysValIleGlyAsnGlySerPheGlyVal 79  
DB 3 CCCTCCCAAGAGTGGCTTACACGACATCAAAAGTGATTGGCAATGGCTCAITTTGGGTC 62  
QY 80 ValTyrGlnAlaLysLeuCysAspSerGlyGluLeuValAlaIleLysLysValLeuGln 99  
DB 63 GTGTACCAAGCAGCGCTGGCAGACACGAGAACTAGTCCCATCAGAAGGTTCTCCAG 122  
QY 100 AspLysArgPheLysAsnArgGluLeuGlnIleMetArgLysLeuAspHisCysAsnIle 119  
DB 123 GACAAGAGGTTCAAGAACCCGAGAGCTGCAGATCATGGCTAAGCTGGACCACTGCAATATT 182  
QY 120 ValArgLeuArgTyrPhePheTyrSerSerGlyGluLysLysAspGluValTyrLeuAsn 139  
DB 183 GTGAGGCTGAGATACTTTTCTACTCCAGTGGCGAGAAGAAAGACGAGCTTTACCTAAAT 242  
QY 140 LeuValLeuAspTyrValProGluThrValTyrArgValAlaArgHisTyrSerArgAla 159  
DB 243 CTGGTGCTGGAATATGTCCCGAGACAGTGTACCGGGTGGCCCGCCACTTCACCAAGGCC 302  
QY 160 LysGlnThrLeuProValIleTyrValLysLeuTyrMetTyrGlnLeuPheArgSerLeu 179  
DB 303 AAGTTGACCATCTCTCTATGTCAAGGTGTACATGTACCATCTCTCCGACGTG 362  
QY 180 AlaTyrIleHisSerPheGlyIleCysHisArgAspIleLysProGlnAsnLeuLeuLeu 199  
DB 363 GCCTACATCCACTCCCGGGGGTGTGTACCGGACATCAGCCCGACACCTGCTGTG 422  
QY 200 AspProAspThrAlaValLeuLysLeuCysAspPheGlySerAlaLysGlnLeuValArg 219  
DB 423 GACCCTGACACTGCTGCTCTCAAGCTCTCCGATTTGGCAGTGCAAGCAGTTGGTCCGA 482

QY 220 GlyGluProAsnValSerTyrIleCysSerArgTyrTyrArgAlaProGluLeuIlePhe 239  
DB 483 GGGGAGGCCAAATGTCTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 542  
QY 240 GlyAlaThrAspTyrThrSerSerIleAspValTrpSerAlaGlyCysValLeuAlaGlu 259  
DB 543 GGAGCCACTGATTACACCTCATCCATCGATGTTGGTCAGCTGGCTGTGTACTGGCAGAG 602  
QY 260 LeuLeuLeuGlyGlnProIlePheProGlyAspSerGlyValAspGlnLeuValGluIle 279  
DB 603 CTCTCTTTGGGCCAGCCCATCTTCCCTGGGACAGTGGGTGGACACAGCTGGTGAGATC 562  
QY 280 IleLysValLeuGlyThrProThrArgGluGlnIleArgGluMetAsnProAsnTyrThr 299  
DB 663 ATCAAGGTGCTGGGAACACCAACCCGGGAACAAATCCGAGACATGAACCCCAACTACAG 722  
QY 300 GluPheLysPheProGlnIleLysAlaHisProTrp-ThrLysValPheArgProArgTh 319  
DB 723 GAGTTCAAGTTCCCTCAGATTANAGCTCACCCCTGGGACAAAGGTGTTCAATCTCGAAC 782  
QY 319 rProProGluAlaIleAlaLeuCysSerArgLysLeuGluTyrThrProThrAlaArgLe 339  
DB 783 GCGGCGACAGGGCATCGCGCTCTGCTCTAGCCTGTGTGAGTACACCCCATCTCTCAGGCT 842  
QY 339 uThrProLeuGluAlaCysAlaHis-SerPhePheAsp 351  
DB 843 CTCCCCACTAGAAGGCTGTGGCAGACAGGCTTCTTTGAT 880

Search completed: September 18, 2006, 03:35:36  
Job time : 4638 secs